# Vector Quantization Codebook Design for Multiuser MIMO

Malcolm Egan

Faculty of Electrical Engineering, Czech Technical University in Prague

*malcolm.egan@agents.fel.cvut.cz*

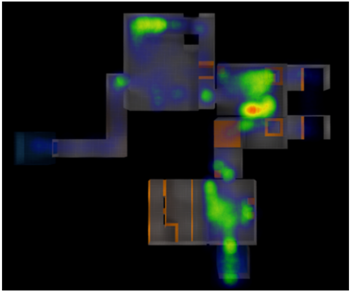
July 14, 2015

# Outline of the Talk

1. Side information in the wireless cellular downlink.
2. What is multiuser MIMO?
3. Side information design for MU-MIMO.
4. Understanding the effect of limited feedback in MU-MIMO.
5. Structured vector quantizers for MU-MIMO with limited feedback.

# Side Information Networks

- Collection of measurements/observations in order to make a decision is a fundamental component of many systems.
- Example: a network of wireless heat sensors measuring the temperature of a room.



# Side Information Network Design

- Data collection is not usually the point of the system $\Rightarrow$ it is *side information*.
- How the measurements are collected, compressed and shared is not optimized for its own sake.
- Instead, the goal is to improve the decision making capabilities of the system.
- Example: the network of wireless heat sensors is designed to control the temperature of the room via a thermostat.
  - The level of accuracy in room temperature required determines the number and quality of the sensors.
  - The level of accuracy required also affects the design of the wireless links.

# The Wireless Cellular Downlink

- In this talk, we are interested in the downlink of wireless cellular networks.

# Side Information in the Wireless Cellular Downlink

- The goal of the downlink is to provide users with the data they have requested.
  - E.g., calls, SMS, or internet data.
- Side information plays an important role in optimizing the transmission from base stations (transmitters) to users (receivers).

# Side Information in the Wireless Cellular Downlink

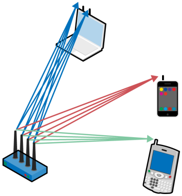
- The side information consists of observations, including:
  1. channel state (e.g., signal attenuation)
  2. queue state (e.g., number of packets in the queue, or the packet delay)
- The side information is often also shared via wireless links.
- The wireless side information links are fundamentally different to the main data link between base stations and users.
  - $\Rightarrow$ not all wireless links in the network have the same purpose.

# Side Information in the Wireless Cellular Downlink

*The side information network is designed to optimize the main data transmissions from base stations to users.*

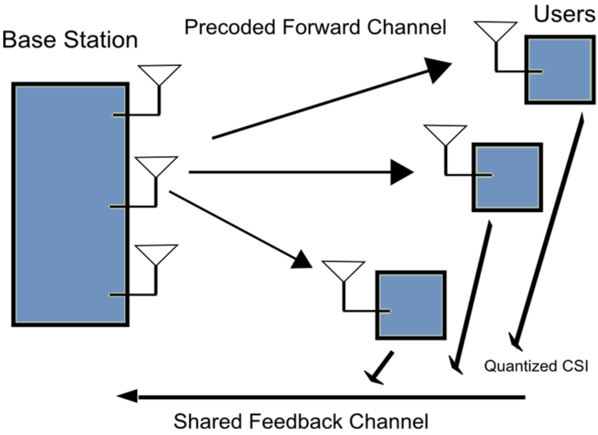
# What is Multiuser MIMO?

- In modern wireless cellular networks, multiple antennas play an important role.
- Multiuser multiple-input multiple-output (MU-MIMO) plays a key role.



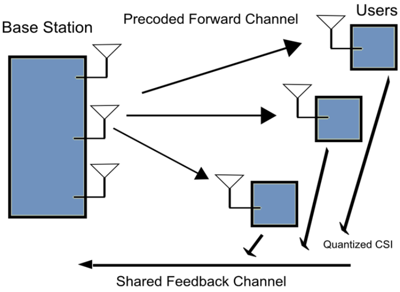
- Why?
  1. Is a key component of recent standards (e.g., LTE, LTE-A).
  2. Forms the basis for advanced interference mitigation techniques (e.g., CoMP, network MIMO).

# The MU-MIMO System Model



# The MU-MIMO System Model

- The MU-MIMO system consists of three basic components:
  1. a $N_t$ antenna base station;
  2. $N_U$ users, each with a single antenna;
  3. a $N_U \times N_t$ channel matrix $\mathbf{H}$.



# The MU-MIMO System Model

- The base station seeks to transmit a data vector $\mathbf{u} \in \mathbb{C}^{N_U}$
  - Each element of $\mathbf{u}$ is the data for a user.
  - The data vector $\mathbf{u}$ is assumed to be Gaussian ($\mathbf{u} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$).
- The base station can perform linear data processing.
- The received vector is then

$$\mathbf{y} = \mathbf{HPVu} + \mathbf{n}, \tag{1}$$

where

1. $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_{N_U}]$ is the linear data processing at the base station;
2. $\mathbf{P} = \mathrm{diag}(p_1/\|\mathbf{v}_1\|^2, \ldots, p_{N_t}/\|\mathbf{v}_{N_t}\|^2)$ of transmit powers for each antenna.
3. $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_{N_U}]^T$ is the channel matrix;
4. $\mathbf{n}$ is Gaussian noise ($\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$)

# The Role of Side Information in MU-MIMO

- Without linear processing, there is inter-user interference.
- Suppose $\mathbf{V} = \mathbf{I}$. This means that

$$y_i = h_{i,i}u_i + \sum_{j \neq i} h_j u_j + n_i. \quad (2)$$

- To mitigate the interference, we can exploit side information.
  - Known as *channel state information at the transmitter*.
- For instance, suppose the channel matrix $\mathbf{H}$ is known to the base station.
- Choose $\mathbf{V} = \mathbf{H}^{-1}$, which means that

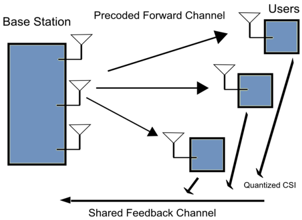
$$y_i = u_i + n_i. \quad (3)$$

# The Role of Side Information in MU-MIMO

- There are two problems with using this approach:
  1. There is a power constraint, which means that

$$\sum_{i=1}^{N_t} p_i \leq P_{\max}. \quad (4)$$

  2. We need to obtain the channel matrix $\mathbf{H}$ at the transmitter.



# MU-MIMO with Limited Feedback

- We can only obtain a quantized (compressed) version of $\mathbf{H}$ at the base station.
- This is known as MU-MIMO with *limited feedback*.
- That is, we have the quantized channel matrix $\hat{\mathbf{H}}$.
- The received signal is then

$$\mathbf{y} = \mathbf{H}\mathbf{P}\hat{\mathbf{H}}^{-1}\mathbf{u} + \mathbf{n}. \tag{5}$$

# The Design Objective

- The goal of the system is to maximize the average total data rate sent to all users.
- By Shannon's theorem for the Gaussian channel, the total data rate for a given $\mathbf{H}$ is given by

$$R = \sum_{i=1}^{N_U} \log\left(1 + \frac{|\mathbf{h}_i^\dagger \mathbf{v}_i|^2 \frac{p_i}{\|\mathbf{v}_i\|^2}}{\sigma^2 + \sum_{j=1, j\neq i}^{n} |\mathbf{h}_i^\dagger \mathbf{v}_j|^2 \frac{p_j}{\|\mathbf{v}_j\|^2}}\right), \quad (6)$$

where the power levels $p_i$ are chosen to satisfy

$$\sum_{i=1}^{N_t} p_i \leq P_{\max}. \quad (7)$$

# The Design Objective

Our problem is to design the quantization scheme to determine $\hat{\mathbf{H}}$ so that the expected rate $\mathbb{E}[R]$ is maximized.

- This involves determining how to quantize each user's channel vector $\mathbf{h}_i^{\dagger}$.

# The Basic Approach

- To quantize each user's channel vector $\mathbf{h}_i^\dagger$, we consider two parts:
  1. the *channel gain*, $\|\mathbf{h}_i^\dagger\|$;
  2. and the *channel shape*

$$\tilde{\mathbf{h}}_i^\dagger = \frac{\mathbf{h}_i^\dagger}{\|\mathbf{h}_i^\dagger\|}. \tag{8}$$

- Quantizing the channel gain is a *scalar quantization* problem.
  - This only requires 3 bits to obtain near optimal performance.
- Quantizing the channel shape is a *vector quantization* problem.
  - This is our focus.

# Quantizing the Channel Shape

- Suppose we have a codebook of vectors $\mathcal{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_N]$, $\mathbf{f}_i \in \mathbb{S}^{N_t - 1}$.
  - I.e., codewords lie on the unit sphere in $N_t$ dimensions.
  - This space can be identified as the complex projective space $\mathbb{C}P^{N_t - 1}$ or Grassmannian manifold $\mathcal{G}(N_t, 1)$.
- Now, suppose user $i$'s channel shape is $\tilde{\mathbf{h}}_i^\dagger$.
- We can quantize the channel shape via

$$k_i^* = \arg \max_{k \in 1,2,\ldots,N} |\tilde{\mathbf{h}}_i^\dagger \mathbf{f}_k|^2. \tag{9}$$

# Design Problems

**There are two key design problems:**

1. How big should the codebook be (i.e., $N$)?
2. How do we construct the codebook?

# How Big Should the Codebook Be?

- This question depends on two factors:
  1. how the codebook is constructed;
  2. and how the power levels are chosen.

# How Big Should the Codebook Be?

- One approach is to make the following assumptions:
  1. the channel shape $\tilde{\mathbf{h}}_i$ is drawn from an isotropic distribution on the unit sphere
  2. the codebook is drawn from an isotropic distribution on the unit sphere;
  3. and the power is equal for each antenna, which means that

$$p_i = \frac{P_{\max}}{N_t}. \tag{10}$$

- This case was studied in [Jindal2006], where it was shown that choosing

$$N = 2^{(N_t - 1)\log_2 P_{\max}} \tag{11}$$

ensures that the rate loss (relative to perfect CSIT) is upper bounded by $N_t$ bits/s/Hz.

# The Case of Optimal Power Control

- It is also possible to gain insight into the effect of quantization using optimal power control, when the quantization error is specified.
- That is, power is allocated by

$$
\begin{aligned}
R^*(\mathrm{vec}(\hat{\mathbf{V}})) = & \max_{p_1,\ldots,p_n} \\
& \sum_{i=1}^{n} \log\left(1 + \frac{|\mathbf{h}_i^\dagger \mathbf{v}_i|^2 \frac{p_i}{\|\mathbf{v}_i\|^2}}{\sigma^2 + \sum_{j=1, j\neq i}^{n} |\mathbf{h}_i^\dagger \mathbf{v}_j|^2 \frac{p_j}{\|\mathbf{v}_j\|^2}}\right) \\
\text{subject to} \quad & \sum_{i=1}^{n} p_i \leq p_{\max} \\
& p_i \geq 0, \quad i = 1, \ldots, n,
\end{aligned}
$$

- We can use techniques from optimization problems with perturbations.
  - Note that the perturbations are in $\mathbb{C}^{N_t N_u}$.

# The Case of Optimal Power Control

- Observe that we can write the Taylor series expansion

$$R^*(\mathrm{vec}(\hat{\mathbf{W}})) = R^*(\mathrm{vec}(\mathbf{W})) + D_{\tilde{\mathbf{d}}} R^*(\mathrm{vec}(\mathbf{W}))\|\mathbf{d}\| + o(\|\mathbf{d}\|^2), \quad (12)$$

where the directional derivative is given by

$$D_{\mathbf{e}} R^*(\mathrm{vec}(\mathbf{W})) = \lim_{t \to 0} \frac{R^*(\mathrm{vec}(\mathbf{W}) + t\tilde{\mathbf{d}}) - R^*(\mathrm{vec}(\mathbf{W}))}{t}. \quad (13)$$

- This can be used to give a rate loss bound

$$|R^*(\mathrm{vec}(\hat{\mathbf{W}})) - R^*(\mathrm{vec}(\mathbf{W}))| \leq |D_{\tilde{\mathbf{d}}} R^*(\mathrm{vec}(\mathbf{W}))|\|\mathbf{d}\| + |o(\|\mathbf{d}\|^2)|, \quad (14)$$

# The Case of Optimal Power Control

- The key difficulty is to obtain the directional derivative $D_{\tilde{\mathbf{d}}} R^*(\mathrm{vec}(\mathbf{W}))$.
- To obtain this we have proved a Danskin-type theorem.
- A key feature of our theorem is that it is in terms of Wirtinger derivatives:

$$\frac{\partial}{\partial z_i} f(z_0) = \frac{1}{2}\left(\frac{\partial}{\partial x_i} f(z_0) - i\frac{\partial}{\partial y_i} f(z_0)\right)$$
$$\frac{\partial}{\partial z_i^*} f(z_0) = \frac{1}{2}\left(\frac{\partial}{\partial x_i} f(z_0) + i\frac{\partial}{\partial y_i} f(z_0)\right),$$

with

$$\frac{\partial f}{\partial \mathbf{z}} = \left[\frac{\partial f}{\partial z_1}, \ldots, \frac{\partial f}{\partial z_n}\right],$$
$$\frac{\partial f}{\partial \mathbf{z}^*} = \left[\frac{\partial f}{\partial z_1^*}, \ldots, \frac{\partial f}{\partial z_n^*}\right]. \quad (15)$$

# The Case of Optimal Power Control

## Theorem

*Let* $m, n, k \in \mathbb{N}$, $f : \mathbb{R}^m \times \mathbb{C}^{nk} \to \mathbb{R}$ *(*$(\mathbf{x}, \mathbf{h}) \mapsto f$*) be a continuous function in* $\mathbb{R}^m \times \mathbb{C}^{nk}$ *and* $\mathcal{B}$ *be a compact subset of* $\mathbb{R}^m$*. Denote* $S(\mathbf{h}_0)$ *as the set of* $\mathbf{x} \in \mathcal{B}$ *that minimizes* $f(\mathbf{x}, \mathbf{h}_0)$ *with* $\mathbf{h}_0$ *fixed. Suppose that* $D_{\mathbf{h}} f(\mathbf{x}, \mathbf{h})$ *exists and is continuous in* $\mathbb{R}^m \times \mathbb{C}^{nk}$*. Then, the optimal value function*

$$v(\mathbf{h}) = \inf_{\mathbf{x} \in \mathcal{B}} f(\mathbf{x}, \mathbf{h}) \tag{16}$$

*is directionally differentiable in the direction* $\tilde{\mathbf{d}} \in \mathbb{C}^{nk}$, $\|\tilde{\mathbf{d}}\| = 1$, *and*

$$D_{\tilde{\mathbf{d}}} v(\mathbf{h}_0) = \min_{\mathbf{x} \in S(\mathbf{h}_0)} \frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}, \mathbf{h}_0)\tilde{\mathbf{d}} + \frac{\partial f}{\partial \mathbf{h}^*}(\mathbf{x}, \mathbf{h}_0)\tilde{\mathbf{d}}^*. \tag{17}$$

# The Case of Optimal Power Control

- Observe that in the case of perfect CSI, the power control problem reduces to

$$\begin{aligned} R^*(\mathrm{vec}(\mathbf{W})) = \max_{p_1,\ldots,p_n} \quad & \sum_{i=1}^{n} \log\left(1 + \frac{p_i}{\sigma^2 \|\mathbf{w}_i\|^2})\right) \\ \text{subject to} \quad & \sum_{i=1}^{n} p_i \leq p_{\max} \\ & p_i \geq 0, \quad i = 1, \ldots, n. \end{aligned} \tag{18}$$

- This problem is convex $\Rightarrow$ we only need to compute the directional derivative at the optimal solution

$$p_{i,opt} = \left(\mu - \sigma^2 \|\mathbf{w}_i\|^2\right)^+, \tag{19}$$

so that $\sum_{i=1}^{n} \left(\mu - \sigma^2 \|\mathbf{w}_i\|^2\right)^+ \leq p_{\max}$ and $\mu \geq 0$ is satisfied.

# The Case of Optimal Power Control

- From these results, we can compute the rate loss bound

$$|R^*(\mathrm{vec}(\hat{\mathbf{W}})) - R^*(\mathrm{vec}(\mathbf{W}))| \leq |D_{\tilde{\mathbf{d}}} R^*(\mathrm{vec}(\mathbf{W}))| \|\mathbf{d}\| + |o(\|\mathbf{d}\|^2)|, \quad (20)$$

- A key observation is that for a sufficiently small perturbation $\|\mathbf{d}\|$, the rate loss scales with $\|\mathbf{d}\|$.

# How Do We Construct the Codebook?

- So far, we have addressed the first key design problem:
  - How big should the codebook be?
- We now turn to the second key design problem.

> How do we construct the quantization codebook $\mathcal{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_N]$?

# Codebook Design Criteria

- The goal of codebook design (i.e., $\mathcal{F}$) is to maximize the average data rate.
- This is difficult, so we resort to optimizing bounds.
- An important design criterion is the *Grassmannian criterion*:

$$\max_{\mathcal{F}} \min_{i \neq j} 1 - |\mathbf{f}_i^{\dagger}\mathbf{f}_j|^2.$$

- This maximizes the distance between codewords in the codebook $\mathcal{F}$.
- An alternative approach is to minimize the *expected square correlation*:

$$\min_{\mathcal{F}} \sum_{i=1}^{N} \sum_{j=1}^{N} |\mathbf{f}_i^{\dagger}\mathbf{f}_j|^2 \mu(\mathbf{f}_i)\mu(\mathbf{f}_j),$$

where $\mu(\cdot)$ is the Haar measure on $\mathcal{U}(N_t)$.

# How Do We Optimize the Design Criteria?

- To optimize the design criteria, we need to optimize over codebooks (i.e., sets of vectors in $\mathbb{C}P^{N_t-1}$.
- In fact, there is more structure: *codebooks are frames*.

## Definition

A sequence $\Phi = (\phi_i)_{i=1,\dots,N} \in \mathbb{C}^{N_t}$ is a frame for $\mathbb{C}^{N_t}$, if there exists constants $0 < A \leq B < \infty$ such that

$$A\|\mathbf{x}\|^2 \leq \sum_{i=1}^{N} |\langle \mathbf{x}, \phi_i \rangle|^2 \leq B\|\mathbf{x}\|^2,$$

for all $\mathbf{x} \in \mathbb{C}^{N_t}$. A frame is *tight* if $A = B$ and unit norm if $\|\phi_i\| = 1$ for all $i$.

# Optimizing the Grassmannian Criterion

- A key result in frame theory is that the Grassmannian criterion

$$\max_{\mathcal{F}} \min_{i \neq j} 1 - |\mathbf{f}_i^\dagger \mathbf{f}_j|^2.$$

is optimized by the *equiangular tight frames*, where

$$|\mathbf{f}_i^\dagger \mathbf{f}_j|^2 = \frac{N - N_t}{N_t(N-1)},\ \ i \neq j.$$

- However, ETFs only exist for $N \leq N_t^2 \Rightarrow$ not suitable for codebooks with large $N$.
- For $N > N_t^2$, we need an alternative approach.

# Optimizing the ESC

- Constructing codebooks that optimize the Grassmannian criterion is difficult.
- An alternative approach is to minimize the ESC

$$ESC = \sum_{i=1}^{N} \sum_{j=1}^{N} |\mathbf{f}_i^{\dagger} \mathbf{f}_j|^2 \mu(\mathbf{f}_i) \mu(\mathbf{f}_j),$$

which has more (easier to find) solutions.

# ESC Optimal Frames

## Theorem

*A locally optimal solution to the ESC minimization is the frame $\Phi$ satisfying:*

1. *the probability of selecting the i-th codeword is given by $\mu(\mathbf{f}_i) = \frac{1}{N}$; i.e., all codewords are equally likely;*
2. *the frame $\Phi$ is tight.*

# ESC Optimal Frames

- A corollary is that for a locally optimal codebook

$$ESC = \frac{1}{N_t}.$$

- This condition is achieved when the conditions in the theorem holds. That is,
  1. $$\mu(\mathbf{f}_i) = \frac{1}{N};$$
  2. the codebook $\mathcal{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_N]$ is tight; i.e.,

$$\mathcal{F}\mathcal{F}^\dagger = \frac{N}{N_t}\mathbf{I}.$$

- It is also easy to check numerically.

# How Do We Construct ESC Optimal Frames

- An important class of frames optimizes the ESC.

**Theorem**

*If $U_1, \ldots, U_N$ is an algebraic group of unitary matrices (with matrix multiplication as the operation), $\phi \in \mathbb{C}^{N_t}$, $\|\phi\| = 1$, and the frame $\Phi$ with elements*

$$[U_1\phi, \ldots, U_N\phi]$$

*is a tight frame. Then, $\Phi$ minimizes the ESC.*

# An Explicit Construction

- Consider the group $SL_2(\mathbb{F}_3)$.
- This has a unitary representation with generator matrices

$$A = \frac{e^{10\pi i/8}}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -1 \end{pmatrix},$$

$$P = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

$$Q = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

# An Explicit Construction

- Construct the group as

$$A^s P^i Q^p, \;\; s = 0, 1, 2, \;\; i = 0, 1, 2, 3, \;\; p = 0, 1.$$

- The choice of

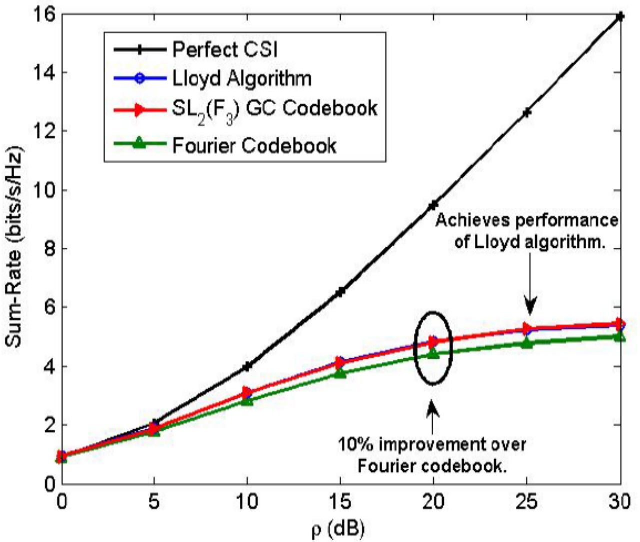
$$\phi = \frac{1}{\sqrt{2}}[1, 1]^T$$

  yields a tight frame.
- We can then obtain a codebook as

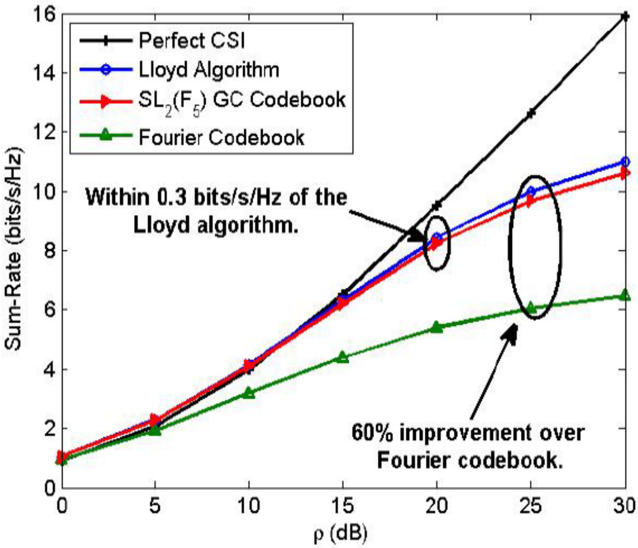
$$\mathcal{F} = \frac{1}{\sqrt{2}}\left[[1, -1]^T, [-1, -1]^T, [0, 1 - i]^T, [1 + i, 0]^T, [-1, i]^T, [-i, 1]^T\right]$$

- This codebook minimizes the ESC.

# A Performance Comparison: $N_t = N_U = 2$, 6 codewords



# A Performance Comparison: $N_t = N_U = 2$, 60 codewords



# Conclusions

- We have considered the problem of designing how to provide side information in MU-MIMO.
- The basic problem is to construct vector quantization codebooks.
- There are two subproblems:
  - How big should the codebook be?
  - How to construct the codebook?
- We proposed a codebook construction technique based on group representation theory.
- Our construction can perform close to optimal.